

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 09/807,990A  
Source: IFW/6  
Date Processed by STIC: 1/13/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:50

Input Set : A:\206523US0PCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

3 <110> APPLICANT: ISHIKAWA, KOHKI  
 4 SUZUKI, EI-ICHIRO  
 5 GONDOH, KEIKO  
 6 SHIMBA, NOBUHISA  
 7 MIHARA, YASUHIRO  
 8 KAWASAKI, HISASHI  
 9 KURAHASHI, OSAMU  
 10 KOUDA, TOHRU  
 11 SHIMAOKA, MEGUMI  
 12 KOZUTSUMI, RIE  
 13 ASANO, YASUHISA  
 15 <120> TITLE OF INVENTION: MUTANT NUCLEOSIDE-5'-PHOSPHATE PRODUCING ENZYMES  
 17 <130> FILE REFERENCE: 206523US0PCT  
 19 <140> CURRENT APPLICATION NUMBER: 09/807,990A  
 C--> 20 <141> CURRENT FILING DATE: 2001-05-03  
 22 <150> PRIOR APPLICATION NUMBER: JP 11/249545  
 23 <151> PRIOR FILING DATE: 1999-09-03  
 25 <160> NUMBER OF SEQ ID NOS: 125  
 27 <170> SOFTWARE: PatentIn version 3.3  
 29 <210> SEQ ID NO: 1  
 30 <211> LENGTH: 1225  
 31 <212> TYPE: DNA  
 32 <213> ORGANISM: Escherichia blattae  
 35 <220> FEATURE:  
 36 <221> NAME/KEY: CDS  
 37 <222> LOCATION: (331)..(1077)  
 39 <400> SEQUENCE: 1

40	ctgcaggcga aaggcaatgt ggtggccggt gagacggcac tctacgagat taaggataag	60
42	taactatcca ttattacagg taacagcatt gtcctgagt gtgatgtcat acctgagcgg	120
44	cgcggggggt ccccgggcgg ctttttttta tggggctgcg gtgaggagcg ttatctgctg	180
46	gccctgtttg tgcaacaaac gcttttattg tgtaattttt gtgacgtata tcagggtttt	240
48	aagcaccctg tggcgctcat actggcaacc tgttgatatt aagcaacact cttcactcac	300
50	ggaattaaca cgcacagtaa aggtatacgc atg aaa aaa cgt gtt ctg gca gtt	354
51	Met Lys Lys Arg Val Leu Ala Val	
52	1 5	
54	tgt ttt gcc gca ttg ttc tct tct cag gcc ctg gcg ctg gtc gct acc	402
55	Cys Phe Ala Ala Leu Phe Ser Ser Gln Ala Leu Ala Leu Val Ala Thr	
56	10 15 20	
58	ggc aac gac act acc acg aaa ccg gat ctc tac tac ctc aag aac agt	450
59	Gly Asn Asp Thr Thr Thr Lys Pro Asp Leu Tyr Tyr Leu Lys Asn Ser	
60	25 30 35 40	
62	gaa gcc att aac agc ctg gcg ctg ttg ccg cca cca ccg gcg gtg ggc	498
63	Glu Ala Ile Asn Ser Leu Ala Leu Leu Pro Pro Pro Pro Ala Val Gly	

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:50

Input Set : A:\206523US0PCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

```

64          45          50          55
66 tcc att gcg ttt ctc aac gat cag gcc atg tat gaa cag ggg cgc ctg      546
67 Ser Ile Ala Phe Leu Asn Asp Gln Ala Met Tyr Glu Gln Gly Arg Leu
68          60          65          70
70 ctg cgc aac acc gaa cgc ggt aag ctg gcg gcg gaa gat gca aac ctg      594
71 Leu Arg Asn Thr Glu Arg Gly Lys Leu Ala Ala Glu Asp Ala Asn Leu
72          75          80          85
74 agc agt ggc ggg gtg gcg aat gct ttc tcc ggc gcg ttt ggt agc ccg      642
75 Ser Ser Gly Gly Val Ala Asn Ala Phe Ser Gly Ala Phe Gly Ser Pro
76          90          95          100
78 atc acc gaa aaa gac gcc ccg gcg ctg cat aaa tta ctg acc aat atg      690
79 Ile Thr Glu Lys Asp Ala Pro Ala Leu His Lys Leu Leu Thr Asn Met
80 105          110          115          120
82 att gag gac gcc ggg gat ctg gcg acc cgc agc gcg aaa gat cac tat      738
83 Ile Glu Asp Ala Gly Asp Leu Ala Thr Arg Ser Ala Lys Asp His Tyr
84          125          130          135
86 atg cgc att cgt ccg ttc gcg ttt tat ggg gtc tct acc tgt aat acc      786
87 Met Arg Ile Arg Pro Phe Ala Phe Tyr Gly Val Ser Thr Cys Asn Thr
88          140          145          150
90 acc gag cag gac aaa ctg tcc aaa aat ggc tct tat ccg tcc ggg cat      834
91 Thr Glu Gln Asp Lys Leu Ser Lys Asn Gly Ser Tyr Pro Ser Gly His
92          155          160          165
94 acc tct atc ggc tgg gct act gcg ctg gtg ctg gca gag atc aac cct      882
95 Thr Ser Ile Gly Trp Ala Thr Ala Leu Val Leu Ala Glu Ile Asn Pro
96          170          175          180
98 cag cgc cag aac gag atc ctg aaa cgc ggt tat gag ctg ggc cag agc      930
99 Gln Arg Gln Asn Glu Ile Leu Lys Arg Gly Tyr Glu Leu Gly Gln Ser
100 185          190          195          200
102 cgg gtg att tgc ggc tac cac tgg cag agt gat gtg gat gcc gcg cgg      978
103 Arg Val Ile Cys Gly Tyr His Trp Gln Ser Asp Val Asp Ala Ala Arg
104          205          210          215
106 gta gtg gga tct gcc gtt gtg gcg acc ctg cat acc aac ccg gcg ttc      1026
107 Val Val Gly Ser Ala Val Val Ala Thr Leu His Thr Asn Pro Ala Phe
108          220          225          230
110 cag cag cag ttg cag aaa gcg aag gcc gaa ttc gcc cag cat cag aag      1074
111 Gln Gln Gln Leu Gln Lys Ala Lys Ala Glu Phe Ala Gln His Gln Lys
112          235          240          245
114 aaa taatcctgac taccgccttg ccttgacagg cggtagtggt ttccactggc      1127
115 Lys
118 cccgattcgc tattcccaca gtaataatga cggtatatga ttttgtgcaa cgaaaagggt      1187
120 gtgtcacgcc acagcttata agatcatgtg ccgttaac      1225
123 <210> SEQ ID NO: 2
124 <211> LENGTH: 249
125 <212> TYPE: PRT
126 <213> ORGANISM: Escherichia blattae
128 <400> SEQUENCE: 2
130 Met Lys Lys Arg Val Leu Ala Val Cys Phe Ala Ala Leu Phe Ser Ser
131 1          5          10          15
134 Gln Ala Leu Ala Leu Val Ala Thr Gly Asn Asp Thr Thr Thr Lys Pro

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:50

Input Set : A:\206523US0PCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

```

135          20          25          30
138 Asp Leu Tyr Tyr Leu Lys Asn Ser Glu Ala Ile Asn Ser Leu Ala Leu
139          35          40          45
142 Leu Pro Pro Pro Pro Ala Val Gly Ser Ile Ala Phe Leu Asn Asp Gln
143          50          55          60
146 Ala Met Tyr Glu Gln Gly Arg Leu Leu Arg Asn Thr Glu Arg Gly Lys
147 65          70          75          80
150 Leu Ala Ala Glu Asp Ala Asn Leu Ser Ser Gly Gly Val Ala Asn Ala
151          85          90          95
154 Phe Ser Gly Ala Phe Gly Ser Pro Ile Thr Glu Lys Asp Ala Pro Ala
155          100          105          110
158 Leu His Lys Leu Leu Thr Asn Met Ile Glu Asp Ala Gly Asp Leu Ala
159          115          120          125
162 Thr Arg Ser Ala Lys Asp His Tyr Met Arg Ile Arg Pro Phe Ala Phe
163          130          135          140
166 Tyr Gly Val Ser Thr Cys Asn Thr Thr Glu Gln Asp Lys Leu Ser Lys
167 145          150          155          160
170 Asn Gly Ser Tyr Pro Ser Gly His Thr Ser Ile Gly Trp Ala Thr Ala
171          165          170          175
174 Leu Val Leu Ala Glu Ile Asn Pro Gln Arg Gln Asn Glu Ile Leu Lys
175          180          185          190
178 Arg Gly Tyr Glu Leu Gly Gln Ser Arg Val Ile Cys Gly Tyr His Trp
179          195          200          205
182 Gln Ser Asp Val Asp Ala Ala Arg Val Val Gly Ser Ala Val Val Ala
183          210          215          220
186 Thr Leu His Thr Asn Pro Ala Phe Gln Gln Gln Leu Gln Lys Ala Lys
187 225          230          235          240
190 Ala Glu Phe Ala Gln His Gln Lys Lys
191          245
194 <210> SEQ ID NO: 3
195 <211> LENGTH: 1344
196 <212> TYPE: DNA
197 <213> ORGANISM: Morganella morganii
200 <220> FEATURE:
201 <221> NAME/KEY: CDS
202 <222> LOCATION: (316)..(1062)
204 <400> SEQUENCE: 3
205 gaattccgga aaatttcatt cattttaatt gttaagaata tgctggcaaa aacaaaaccc 60
207 aatgctttat attttcttat aatatctgtg tgttatcttt ttcaatacta tcggtcaggt 120
209 cttattttatc cgttcgtaa caaaagccat gctgtttctg tcaaattatc tgaaaatcat 180
211 catcaaaaat acttacctgt cttccgtctg ttccgtcaca cttttttgaa agagttaaca 240
213 tcaatttgca tctctccgcc ctacactggc agacagggtt ctgagtaata ctgttgatc 300
215 tgataaggag atgtc atg aag aag aat att atc gcc ggt tgt ctg ttc tca 351
216          Met Lys Lys Asn Ile Ile Ala Gly Cys Leu Phe Ser
217          1          5          10
219 ctg ttt tcc ctt tcc gcg ctg gcc gcg atc ccg gcg ggc aac gat gcc 399
220 Leu Phe Ser Leu Ser Ala Leu Ala Ala Ile Pro Ala Gly Asn Asp Ala
221          15          20          25
223 acc acc aag ccg gat tta tat tat ctg aaa aat gaa cag gct atc gac 447

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:50

Input Set : A:\206523USOPCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

```

224 Thr Thr Lys Pro Asp Leu Tyr Tyr Leu Lys Asn Glu Gln Ala Ile Asp
225      30      35      40
227 agc ctg aaa ctg tta ccg cca ccg ccg gaa gtc ggc agt att cag ttt      495
228 Ser Leu Lys Leu Leu Pro Pro Pro Pro Glu Val Gly Ser Ile Gln Phe
229 45      50      55      60
231 tta aat gat cag gca atg tat gag aaa ggc cgt atg ctg cgc aat acc      543
232 Leu Asn Asp Gln Ala Met Tyr Glu Lys Gly Arg Met Leu Arg Asn Thr
233      65      70      75
235 gag cgc gga aaa cag gca cag gca gat gct gac ctg gcc gca ggg ggt      591
236 Glu Arg Gly Lys Gln Ala Gln Ala Asp Ala Asp Leu Ala Ala Gly Gly
237      80      85      90
239 gtg gca acc gca ttt tca ggg gca ttc ggc tat ccg ata acc gaa aaa      639
240 Val Ala Thr Ala Phe Ser Gly Ala Phe Gly Tyr Pro Ile Thr Glu Lys
241      95      100      105
243 gac tct ccg gag ctg tat aaa ctg ctg acc aat atg att gag gat gcc      687
244 Asp Ser Pro Glu Leu Tyr Lys Leu Leu Thr Asn Met Ile Glu Asp Ala
245      110      115      120
247 ggt gat ctt gcc acc cgc tcc gcc aaa gaa cat tac atg cgc atc cgg      735
248 Gly Asp Leu Ala Thr Arg Ser Ala Lys Glu His Tyr Met Arg Ile Arg
249 125      130      135      140
251 ccg ttt gcg ttt tac ggc aca gaa acc tgt aat acc aaa gat cag aaa      783
252 Pro Phe Ala Phe Tyr Gly Thr Glu Thr Cys Asn Thr Lys Asp Gln Lys
253      145      150      155
255 aaa ctc tcc acc aac gga tct tac ccg tca ggt cat acg tct atc ggc      831
256 Lys Leu Ser Thr Asn Gly Ser Tyr Pro Ser Gly His Thr Ser Ile Gly
257      160      165      170
259 tgg gca acc gca ctg gtg ctg gcg gaa gtg aac ccg gca aat cag gat      879
260 Trp Ala Thr Ala Leu Val Leu Ala Glu Val Asn Pro Ala Asn Gln Asp
261      175      180      185
263 gcg att ctg gaa cgg ggt tat cag ctc gga cag agc cgg gtg att tgc      927
264 Ala Ile Leu Glu Arg Gly Tyr Gln Leu Gly Gln Ser Arg Val Ile Cys
265      190      195      200
267 ggc tat cac tgg cag agt gat gtg gat gcc gcg cgg att gtc ggt tca      975
268 Gly Tyr His Trp Gln Ser Asp Val Asp Ala Ala Arg Ile Val Gly Ser
269 205      210      215      220
271 gcc gct gtc gcg aca tta cat tcc gat ccg gca ttt cag gcg cag tta      1023
272 Ala Ala Val Ala Thr Leu His Ser Asp Pro Ala Phe Gln Ala Gln Leu
273      225      230      235
275 gcg aaa gcc aaa cag gaa ttt gca caa aaa tca cag aaa taaaagcagt      1072
276 Ala Lys Ala Lys Gln Glu Phe Ala Gln Lys Ser Gln Lys
277      240      245
279 gatatctggt cagggcagtg caatatctgc cctgaaatcc ctgtttattc ccacatccag      1132
281 cggctcttccc gatcccagcc ttttggtttc atgcagctgt agaaatagcg gttgcggctg      1192
283 tcttcattca catccatcac ataactttcc gttaccggtg tctgctcttt gtaggttttg      1252
285 ctgttaccgc agtcatcgtc ttttttgtag cgtttctcca catcccgcat cacactgcgc      1312
287 tgagcaactt catttttcac cggataaagc tt      1344
290 <210> SEQ ID NO: 4
291 <211> LENGTH: 249
292 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:50

Input Set : A:\206523US0PCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

```

293 <213> ORGANISM: Morganella morganii
295 <400> SEQUENCE: 4
297 Met Lys Lys Asn Ile Ile Ala Gly Cys Leu Phe Ser Leu Phe Ser Leu
298 1 5 10 15
301 Ser Ala Leu Ala Ala Ile Pro Ala Gly Asn Asp Ala Thr Thr Lys Pro
302 20 25 30
305 Asp Leu Tyr Tyr Leu Lys Asn Glu Gln Ala Ile Asp Ser Leu Lys Leu
306 35 40 45
309 Leu Pro Pro Pro Pro Glu Val Gly Ser Ile Gln Phe Leu Asn Asp Gln
310 50 55 60
313 Ala Met Tyr Glu Lys Gly Arg Met Leu Arg Asn Thr Glu Arg Gly Lys
314 65 70 75 80
317 Gln Ala Gln Ala Asp Ala Asp Leu Ala Ala Gly Gly Val Ala Thr Ala
318 85 90 95
321 Phe Ser Gly Ala Phe Gly Tyr Pro Ile Thr Glu Lys Asp Ser Pro Glu
322 100 105 110
325 Leu Tyr Lys Leu Leu Thr Asn Met Ile Glu Asp Ala Gly Asp Leu Ala
326 115 120 125
329 Thr Arg Ser Ala Lys Glu His Tyr Met Arg Ile Arg Pro Phe Ala Phe
330 130 135 140
333 Tyr Gly Thr Glu Thr Cys Asn Thr Lys Asp Gln Lys Lys Leu Ser Thr
334 145 150 155 160
337 Asn Gly Ser Tyr Pro Ser Gly His Thr Ser Ile Gly Trp Ala Thr Ala
338 165 170 175
341 Leu Val Leu Ala Glu Val Asn Pro Ala Asn Gln Asp Ala Ile Leu Glu
342 180 185 190
345 Arg Gly Tyr Gln Leu Gly Gln Ser Arg Val Ile Cys Gly Tyr His Trp
346 195 200 205
349 Gln Ser Asp Val Asp Ala Ala Arg Ile Val Gly Ser Ala Ala Val Ala
350 210 215 220
353 Thr Leu His Ser Asp Pro Ala Phe Gln Ala Gln Leu Ala Lys Ala Lys
354 225 230 235 240
357 Gln Glu Phe Ala Gln Lys Ser Gln Lys
358 245
361 <210> SEQ ID NO: 5
362 <211> LENGTH: 991
363 <212> TYPE: DNA
364 <213> ORGANISM: Salmonella typhimurium
367 <220> FEATURE:
368 <221> NAME/KEY: CDS
369 <222> LOCATION: (132)..(827)
371 <400> SEQUENCE: 5
372 cagtcggtg tggacagacg ataatgccag ggcgacgcgtc ctgctttttt acctgtatgt 60
374 tgaataacca ttgcaataaa tcattatagg attacatctg tttattattg cctgatccgg 120
376 agtgagtcctt t atg aaa agt cgt tat tta gta ttt ttt cta cca ctg atc 170
377 Met Lys Ser Arg Tyr Leu Val Phe Phe Leu Pro Leu Ile
378 1 5 10
380 gta gct aaa tat aca tca gca gaa aca gtg caa ccc ttt cat tct cct 218
381 Val Ala Lys Tyr Thr Ser Ala Glu Thr Val Gln Pro Phe His Ser Pro

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 01/13/2005  
PATENT APPLICATION:    US/09/807,990A      TIME: 12:07:51

Input Set : A:\206523USOPCT.txt  
Output Set: N:\CRF4\01132005\I807990A.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:121; Xaa Pos. 2,3,4,5,6,7

Seq#:123; Xaa Pos. 3,4,5,6,7,9,10,11

**VERIFICATION SUMMARY**

DATE: 01/13/2005

PATENT APPLICATION: US/09/807,990A

TIME: 12:07:51

Input Set : A:\206523US0PCT.txt

Output Set: N:\CRF4\01132005\I807990A.raw

L:20 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:2292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:121 after pos.:0

L:2331 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123 after pos.:0